

DOCUMENTS DE TRAVAIL - WORKING PAPERS

W.P. 02-03

Market-making, inventories and martingale pricing

Patrick ROGER
Christian AT
Laurent FLOCHEL

Juin 2002

GATE Groupe d'Analyse et de Théorie Économique
UMR 5824 du CNRS
93 chemin des Mouilles – 69130 Écully – France
B.P. 167 – 69131 Écully Cedex
Tél. +33 (0)4 72 86 60 60 – Fax +33 (0)4 72 86 60 90
Messagerie électronique gate@gate.cnrs.fr
Serveur Web : www.gate.cnrs.fr

Market-making, inventories and martingale pricing

Patrick Roger, LARGE, University Louis Pasteur, Strasbourg I

Christian At, GATE, UMR 5824 CNRS, University Lumière Lyon II

Laurent Flochel, GATE, UMR 5824 CNRS, University Lumière Lyon II

Abstract

We discuss Shen and Starr(2002) results and show that the bid-ask spread of a monopolistic market-maker doesn't depend on his inventory when he posts "martingale prices " in an inventory model with random volumes and an unknown direction of trade.

Key words : martingale pricing, market maker, bid ask spread.

Résumé

Nous discutons les résultats de Shen et Starr (2002) et montrons que la fourchette de prix de réservation d'un teneur de marché ne dépend pas de son inventaire lorsque celui-ci fixe un prix suivant une martingale dans un modèle d'inventaire avec volumes aléatoires et direction de l'échange inconnue.

JEL Classification : G12,D42

Mots clés : martingale, teneur de marché, fourchette de prix.

1 Introduction

In a recent paper, Shen and Starr (2002) present an inventory model of market-making based on two crucial assumptions; the first one is that the market-maker's wealth is a martingale and the second one is related the existence of an exogenous price process, the bid and ask prices being symmetric with respect to this exogenous price.

The market-maker (referred to as the MM in the following) then posts bid and ask reservation prices at the beginning of each period, where "reservation prices" means prices which equalize the realized wealth at date t and the expected wealth at date $t + 1$, conditioned on current information. It is then an inventory model in the spirit of Amihud and Mendelson (1980), or Ho and Stoll (1981-1983) in the sense that the inventory level is a key variable to determine the next period bid and ask prices and the spread between the two. The main conclusions are that the spread is positively related to the absolute value of the net position of the market-maker, the volatility of volumes and the volatility of prices.

In this paper, we pursue two objectives ; first, we describe the trading process and the behavior of a monopolistic risk averse MM to illustrate that, in such a trading mechanism, the price process is not exogenous when it is defined as the midpoint between the bid and the ask price, as considered by Shen and Starr. Second, if risk aversion is introduced by means of a concave utility function and if the expected utility (with respect to the terminal payment of the risky asset) of the MM is a martingale, then the bid-ask spread is independent of the inventory level while, at the same time, bid and ask prices are determined by this level of inventory. In other words, we argue that the relationship between spread and inventory, established by Shen and Starr, comes from the rather strong assumption of the observability of an exogenous price process when, in fact, only bid and ask prices are observable, the transaction prices being realized on one of these two bounds. It is worth noticing that, even if trades are executed inside the bid-ask interval, the transaction price is not exogenous but depends on the net position of the MM, his preferences and his expected terminal value of the risky asset.

The relationship between spread and inventory has been studied, in a static framework, by Eeckhoudt-Roger (1999) in the context of expected utility models and by Roger (2000) in the RDEU model. They show that the

expected value of the risky asset (which may be interpreted as the exogenous price of Shen and Starr) is not in the bid-ask interval as soon as the MM holds an inventory. These authors also illustrate a non-monotonic relationship between spread and inventory by considering a risky asset which payment is bounded. However, their model is based on non linear pricing rules, that is to say, the pricing schedule posted by the MM is one-to-one with respect to quantities.

In this paper, we deal with linear prices in the sense that the MM posts unit prices and accepts all orders up to a given maximum quantity. In section 2, we first describe the financial market and then the trading mechanism, including the market-maker behavior. In section 3, we prove the invariance of the bid-ask spread with respect to the inventory; the assumption of a mean-variance utility function, often used in microstructure models (see for example Biais, 1993) allows us to obtain simple analytical formulations for prices and spreads. A brief comparison with the spread obtained in the Ho and Stoll model is provided. Section 4 concludes the paper by suggesting directions for future research.

2 The trading process

2.1 Notations and definitions

We consider a discrete-time model with a relevant set of dates $\mathcal{T} = \{0, \dots, T\}$; the financial market is reduced to a single risky security which generates a square-integrable random terminal payment X with $E(X) = \mu_X$ and $\sigma^2(X) = \sigma_X^2$. The uncertainty is described by a filtered probability space $(\Omega, \mathcal{A}, \mathcal{F}, P)$ where Ω is the set of states of nature and \mathcal{F} is a filtration verifying $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{A}$. Moreover, for reasons which will become clear later, we assume that $\mathcal{F}_t \neq \mathcal{F}_{t+1}$ meaning that the sequence of σ -algebras \mathcal{F} is strictly increasing for inclusion.

As we want to focus on the inventory effect, we suppose that X is \mathcal{F}_T -measurable but independent of $\mathcal{F}_t, t < T$; in other words, nothing is revealed to the market before date T , concerning the terminal payment X . The trading process is driven by a risk-averse MM whose job is only to provide liquidity to other traders. He doesn't bear any operating cost and posts, after date- t transactions, bid and ask prices at which he accepts all orders coming from

the market up to a maximum volume α_M (on each side); these orders are executed at date $t + 1$.

The order flow is formalized as follows; let A_t and B_t the following events :

$$\begin{aligned} A_t &= \{\text{buying order from the market at date } t\} \\ B_t &= \{\text{selling order from the market at date } t\} \end{aligned}$$

We assume equal probabilities for the two events, $P(A_t) = P(B_t) = \lambda$ but different probabilities would not change the nature of our results. When a buying (resp. selling) order arrives at date t , its random volume is denoted as $V_t^A (V_t^B)$. To keep results simple, we assume uniform distributions on $[0; \alpha_M]$ for the variables V_t^A and V_t^B . Moreover, $X, \mathbf{1}_{A_t}, \mathbf{1}_{B_t}, V_t^A, V_t^B$ are assumed to be independent. The independence between $\mathbf{1}_{A_t}$ and V_t^A or between $\mathbf{1}_{B_t}$ and V_t^B , which may seem surprising, will be justified in our definition of reservation prices. Let us now denote $\xi_t = (\mathbf{1}_{A_t}, \mathbf{1}_{B_t}, V_t^A, V_t^B)$; the filtration \mathcal{F} is characterized by :

$$\begin{aligned} \mathcal{F}_t &= \sigma(\xi_s, s \leq t) \text{ if } t < T \\ \mathcal{F}_T &= \mathcal{A} = \sigma(X, \xi_s, s \leq T) \end{aligned}$$

These assumptions mean that the relevant information at intermediate dates only concern the trading volumes; as mentioned before, nothing is revealed about X before the final date. Roughly speaking, it is equivalent to assume a constant exogenous price process because $E(X | \mathcal{F}_t) = E(X)$ for $t < T$.

It is worth noticing that simultaneous trades are allowed, since $P(A_t \cap B_t) = \lambda^2 > 0$. Date- t transactions are realized at prices (π_t^A, π_t^B) , posted by the market-maker at date $t - 1$; in fact, when the MM posts his prices, he doesn't know which volumes are to be traded at these prices. The stochastic process of prices (π^A, π^B) is then predictable with respect to \mathcal{F} .

2.2 The market-maker

The dynamics of the MM's wealth and inventory are described by the following equations :

$$\begin{aligned} W_t &= W_{t-1} + \mathbf{1}_{A_t} V_t^A \pi_t^A - \mathbf{1}_{B_t} V_t^B \pi_t^B \\ I_t &= I_{t-1} - \mathbf{1}_{A_t} V_t^A + \mathbf{1}_{B_t} V_t^B \end{aligned} \tag{1}$$

This formulation clarifies the assumption of independence between volumes and indicator functions because volumes change inventory and wealth only if indicator functions are equal to 1. The relationships (1) also mean that the conditional expected wealth of the MM at date $t + 1$ can be written as :

$$E_t(W_{t+1} + I_{t+1}X) = W_t + I_t\mu_X + E_t(\mathbf{1}_{A_{t+1}}V_{t+1}^A(\pi_{t+1}^A - X) - \mathbf{1}_{B_{t+1}}V_{t+1}^B(\pi_{t+1}^B - X))$$

The MM is initially endowed with a sure wealth W_0 and no risky asset. His expected utility at date t for a random wealth level \widetilde{W} is assumed to be of the following form :

$$E_t \left[U \left(\widetilde{W} \right) \right] = E_t(\widetilde{W}) - a\sigma_t^2(\widetilde{W})$$

where a is a parameter of risk aversion, $E_t(\cdot)$ ($\sigma_t^2(\cdot)$) is a shorthand for $E(\cdot | \mathcal{F}_t)$ ($\sigma^2(\cdot | \mathcal{F}_t)$).

After trades up to date t , the random wealth of the MM is $W_t + I_tX$ and we can write :

$$E_t(U(W_t + I_tX)) = W_t + I_t\mu_X - aI_t^2\sigma_X^2$$

because X is independent of \mathcal{F}_t , $t < T$ and W_t and I_t are \mathcal{F}_t -mesurable.

As the market-maker is assumed to start with a wealth W_0 and no risky asset, his utility at date 0 is equal to W_0 . However, $E_t(U(W_t + I_tX))$ depends on I_t , which is the result of the order flow, not controlled by the MM, but also depends on W_t , which is determined by the MM's pricing policy. In the following section, we analyze the case of "martingale pricing"; it consists, for the market maker, to post prices that equalize the expected utility effectively obtained at date t to the conditional expectation of date $t+1$ utility. In other words, conditional on the history of the trading process up to date t , the MM posts his reservation prices to be indifferent between trading and not trading during the next period.

3 Martingale pricing

3.1 Spread and inventory

The way the MM posts bid and ask reservation prices is described in the following definition.

Definition 1 *At date t , the reservation prices $(\pi_{t+1}^B, \pi_{t+1}^A)$ posted by the MM verify :*

$$E_t(U(W_t + I_t X)) = E_t(U(W_{t+1} + I_{t+1} X))$$

where W_{t+1} and I_{t+1} are given by relation (1)

Our essential result is then the following.

Proposition 2 *At every date t , the spread is equal to :*

$$\pi_t^A - \pi_t^B = a\sigma_X^2 \alpha_M \left[\frac{4}{3} - \lambda \right]$$

and is then independent from I_t .

Proof : Suppose that, at date t , the MM has a net inventory $I_t = \alpha > 0$ and a sure wealth W_t . The two reservation prices for trading at date $t + 1$ can be formulated by means of a function $M_{t+1}^\alpha(v, w)$ defined by :

$$M_{t+1}^\alpha(v, w) = v(\mu_X - \pi_{t+1}^B) - w(\mu_X - \pi_{t+1}^A) - a\sigma_X^2(v + \alpha - w)^2 \quad (2)$$

$W_t + \alpha\mu_X + M_{t+1}^\alpha(v, w)$ is the expected utility after a trade volume v on the bid side and w on the ask side, when the inventory is α .

Recalling the definition of λ , the equalities $P(A \cap B^c) = P(A^c \cap B) = \lambda(1 - \lambda)$ and $P(A \cap B) = \lambda^2$, the assumption about the probability distribution of trading volumes leads to the following definition.

Definition 3 *When the level of inventory is equal to α , the bid and ask prices π_{t+1}^B and π_{t+1}^A are a solution of the following equation:*

$$\begin{aligned} & \frac{1}{\alpha_M^2} \int_0^{\alpha_M} \int_0^{\alpha_M} \lambda(1 - \lambda) [M_{t+1}^\alpha(v, 0) + M_{t+1}^\alpha(0, w)] dv dw \\ & + \frac{1}{\alpha_M^2} \int_0^{\alpha_M} \int_0^{\alpha_M} \lambda^2 M_{t+1}^\alpha(v, w) + (1 - \lambda)^2 M_{t+1}^\alpha(0, 0) dv dw \\ & = E_t[U(W_t + \alpha X)] - W_t - \alpha\mu_X = -a\alpha^2\sigma_X^2 \end{aligned} \quad (3)$$

We can remark that $M_{t+1}^\alpha(0, 0)$ is in fact equal to $E_t[U(W_t + \alpha X) - W_t - \alpha X]$. Consequently, definition 3 can be reformulated in the following way :

$$\begin{aligned} & \int_0^{\alpha_M} \int_0^{\alpha_M} (\lambda(1 - \lambda) [M_{t+1}^\alpha(v, 0) + M_{t+1}^\alpha(0, w)] + \lambda^2 M_{t+1}^\alpha(v, w)) dv dw \\ &= \lambda(2 - \lambda) E_t[U(W_t + \alpha X) - W_t - \alpha X] = -a\alpha^2 \sigma_X^2 \lambda(2 - \lambda) \end{aligned}$$

We can now prove the proposition, by defining the three following components :

$$\begin{aligned} A_v &= \frac{1}{\alpha_M} \int_0^{\alpha_M} \lambda(1 - \lambda) M_{t+1}^\alpha(v, 0) dv \\ A_w &= \frac{1}{\alpha_M} \int_0^{\alpha_M} \lambda(1 - \lambda) M_{t+1}^\alpha(0, w) dw \\ A_{vw} &= \frac{1}{\alpha_M^2} \int_0^{\alpha_M} \int_0^{\alpha_M} \lambda^2 M_{t+1}^\alpha(v, w) dv dw \end{aligned}$$

The computation of $\Phi = A_v + A_w + A_{vw}$ is reported in the appendix and leads to the desired result.

Even if the calculations are tedious, the intuition behind this result is quite simple. Using martingale pricing, the MM doesn't care to previous gains and losses when he posts prices; more precisely, he takes $E_t[U(W_t + \alpha X)]$ as the objective for the next date. Consequently, the inventory effect is neutral because it is yet taken into account in the expected utility realized by date t . However, if bid and ask prices were considered separately, without considering the uncertainty about the direction of the next trade, as in the Ho and Stoll (1983) model, each of the two prices would depend on inventory. This point is briefly illustrated hereafter.

3.2 Comparison with the Ho and Stoll model

For the two models to be comparable, we consider a fixed trading volume equal to $\frac{\alpha_M}{2}$, that is the mean volume of the preceding section. The two prices are then solutions of the following equations :

$$\begin{aligned} -a\alpha^2 \sigma_X^2 &= \frac{\alpha_M}{2} (\mu_X - \pi_{t+1}^B) - a\sigma_X^2 \left(\frac{\alpha_M}{2} + \alpha \right)^2 \\ -a\alpha^2 \sigma_X^2 &= \frac{\alpha_M}{2} (\pi_{t+1}^A - \mu_X) - a\sigma_X^2 \left(\alpha - \frac{\alpha_M}{2} \right)^2 \end{aligned}$$

from which we deduce :

$$\begin{aligned}\pi_{t+1}^B &= \mu_X - a\sigma_X^2 \left(2\alpha + \frac{\alpha_M}{2}\right) \\ \pi_{t+1}^A &= \mu_X + a\sigma_X^2 \left(\frac{\alpha_M}{2} - 2\alpha\right) \\ \pi_{t+1}^A - \pi_{t+1}^B &= a\sigma_X^2 \alpha_M\end{aligned}$$

The spread is independent of the inventory level; it is also different from the one obtained in the preceding section for two reasons. First, as we consider random volumes, the position of the MM is riskier and it explains the coefficient $\frac{4}{3}$. At the same time, simultaneous buy and sell orders are possible with probability λ^2 . They induce a hedging effect which decreases the spread, as this effect appears in the spread by means of the negative coefficient $-\lambda$. In other words, λ is a measure of market liquidity; the more liquid is the market, the greater is the possibility of simultaneous trades, leading to an increase in the MM's wealth.

However, even if $\lambda = 1$, the MM is still in a risky situation because the volumes on each side do not cancel each other, as it is the case in the Ho and Stoll framework with known and equal volumes on each side.

4 Conclusion

In this paper, we have demonstrated, in a simple framework, that the bid-ask spread and the inventory level are independent when the MM uses "martingale pricing"; our result also illustrates that the pricing process, defined as the midpoint between the ask and the bid price, is not exogenous and is clearly linked to the inventory level. The different results obtained by Shen and Starr, linking the spread to the inventory level, come from the assumption that bid and ask prices are symmetrical with respect to an exogenous non observable price process. Our formulation also points out the increase in spread due to random volumes and the decrease due to market liquidity, represented by the probability of simultaneous orders on the two sides.

A more fundamental question relies on the martingale assumption; in fact a MM is more likely to post prices in a mean-reverting way, that is, after a "bad trade", he will probably try to recover quickly by lowering his bid price and increasing his ask price. This situation can be taken into account

by considering that the goal of the MM is always to come back to his initial situation, a shorthand for what is usually called the desired inventory level. In such a situation, the bid-ask spread is not independent of the inventory level; however, the relationship is, in general, not monotonic and depends essentially on the history of trades. It is due to the fact that prices which equalize initial utility and conditional expected utility generate benefits for low trade volumes and losses for large volumes¹. Consequently, when volumes are random, the history of trades is a key factor to relate the bid-ask spread and the inventory level.

References

- [1] Amihud, Y. and H. Mendelson, 1980, Dealership market : market making with inventory, *Journal of Financial Economics* 8, 31-53.
- [2] At, C., Flochel L. and P. Roger, 2002, Reservation prices, risk aversion and inventories : the linear case, *Mimeo*, University Louis Pasteur.
- [3] Biais, B., 1993, Price Formation and equilibrium liquidity in fragmented and centralized markets. *Journal of Finance* 48, 157-185.
- [4] Eeckhoudt, L. and P. Roger, 1999, Risk aversion and the bid-ask spread. *European Financial Management*, 5, 323-340.
- [5] Ho, T. and H. Stoll, 1981, Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics* 9, 47-73
- [6] Ho, T., Stoll, H.R., 1983, "The dynamics of dealer markets under competition", *Journal of Finance*, 38, 1053-1074.
- [7] Roger, P., 2000, "Properties of bid and ask reservation prices in the rank-dependent expected utility model", *Journal of Mathematical Economics*, 34, 269-285.
- [8] Shen, P. and M. Starr, 2002, Market-makers' supply of financial market liquidity, *Economics Letters*, 76, n° 1, 53-58.

¹See At, Flochel, Roger (2002).

Appendix

Proof of proposition 2

1) Evaluation of A_v

$$A_v = \frac{\lambda(1-\lambda)}{\alpha_M} \int_0^{\alpha_M} [v(\mu_X - \pi_B(\alpha)) - a\sigma_X^2(v+\alpha)^2] dv$$

Integration with respect to v gives :

$$\begin{aligned} A_v &= \lambda(1-\lambda) \left[\frac{\alpha_M}{2} (\mu_X - \pi_B(\alpha)) - \frac{a\sigma_X^2}{3\alpha_M} ((\alpha_M + \alpha)^3 - \alpha^3) \right] \\ &= \lambda(1-\lambda) \left[\frac{\alpha_M}{2} (\mu_X - \pi_B(\alpha)) - \frac{a\sigma_X^2}{3} (\alpha_M^2 + 3\alpha^2 + 3\alpha_M\alpha) \right] \end{aligned}$$

2) Evaluation of A_w

$$\begin{aligned} A_w &= \frac{\lambda(1-\lambda)}{\alpha_M} \int_0^{\alpha_M} (-w(\mu_X - \pi_{t+1}^A(\alpha)) - a\sigma_X^2(w-\alpha)^2) dw \\ &= \lambda(1-\lambda) \left[-\frac{\alpha_M}{2} (\mu_X - \pi_{t+1}^A(\alpha)) - \frac{a\sigma_X^2}{3} (\alpha_M^2 + 3\alpha^2 - 3\alpha\alpha_M) \right] \end{aligned}$$

3) Evaluation of $A_{v,w}$

$$\begin{aligned} A_{vw} &= \frac{\lambda^2}{\alpha_M^2} \int_0^{\alpha_M} \int_0^{\alpha_M} [v(\mu_X - \pi_{t+1}^B(\alpha)) - w(\mu_X - \pi_{t+1}^A(\alpha)) - a\sigma_X^2(v+\alpha-w)^2] dv dw \\ &= \frac{\lambda^2}{\alpha_M} \int_0^{\alpha_M} \left[\frac{\alpha_M}{2} (\mu_X - \pi_{t+1}^B(\alpha)) - w(\mu_X - \pi_{t+1}^A(\alpha)) - a\sigma_X^2 \frac{(\alpha_M + \alpha - w)^3 - (\alpha - w)^3}{3\alpha_M} \right] dw \\ &= \lambda^2 \left[\frac{\alpha_M}{2} (\pi_{t+1}^A(\alpha) - \pi_{t+1}^B(\alpha)) - a\sigma_X^2 \left[\frac{-\alpha^4 + (\alpha_M + \alpha)^4 + (\alpha - \alpha_M)^4 - \alpha^4}{12\alpha_M^2} \right] \right] \\ &= \lambda^2 \left[\frac{\alpha_M}{2} (\pi_{t+1}^A(\alpha) - \pi_{t+1}^B(\alpha)) - a\sigma_X^2 \left[\frac{\alpha_M^2 + 6\alpha^2}{6} \right] \right] \end{aligned}$$

We then deduce $\Phi = A_v + A_w + A_{vw}$:

$$\begin{aligned} \Phi &= \lambda(1-\lambda) \left[\frac{\alpha_M}{2} (\mu_X - \pi_{t+1}^B) - \frac{a\sigma_X^2}{3} (\alpha_M^2 + 3\alpha^2 + 3\alpha_M\alpha) \right] + \\ &\quad \lambda(1-\lambda) \left[-\frac{\alpha_M}{2} (\mu_X - \pi_{t+1}^A) - \frac{a\sigma_X^2}{3} (\alpha_M^2 + 3\alpha^2 - 3\alpha\alpha_M) \right] + \\ &\quad \lambda^2 \left[\frac{\alpha_M}{2} (\pi_{t+1}^A - \pi_{t+1}^B) - a\sigma_X^2 \left[\frac{\alpha_M^2 + 6\alpha^2}{6} \right] \right] \end{aligned}$$

$$\begin{aligned}
\Phi &= \frac{\lambda\alpha_M}{2} (\pi_{t+1}^A - \pi_{t+1}^B) - \lambda(1-\lambda) \frac{a\sigma_X^2}{3} (2\alpha_M^2 + 6\alpha^2) - \lambda^2 a\sigma_X^2 \left[\frac{\alpha_M^2 + 6\alpha^2}{6} \right] \\
&= \frac{\lambda\alpha_M}{2} (\pi_{t+1}^A - \pi_{t+1}^B) - \frac{a\sigma_X^2}{3} \left[\lambda(1-\lambda) (2\alpha_M^2 + 6\alpha^2) + \lambda^2 \left[\frac{\alpha_M^2 + 6\alpha^2}{2} \right] \right] \\
&= \frac{\lambda\alpha_M}{2} (\pi_{t+1}^A - \pi_{t+1}^B) - \frac{a\sigma_X^2}{3} \left[\alpha_M^2 \left[2\lambda(1-\lambda) + \frac{\lambda^2}{2} \right] + \alpha^2 [6\lambda(1-\lambda) + 3\lambda^2] \right]
\end{aligned}$$

Finally it remains to solve :

$$-a\alpha^2\sigma_X^2\lambda(2-\lambda) = \Phi$$

which is equivalent to :

$$\pi_{t+1}^A - \pi_{t+1}^B = a\sigma_X^2\alpha_M \left[\frac{4}{3} - \lambda \right]$$